

# ADAPTIVE PSEUDO-TRANSIENT-CONTINUATION-GALERKIN METHODS FOR SEMILINEAR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

MARIO AMREIN AND THOMAS P. WIHLE

**ABSTRACT.** In this paper we investigate the application of pseudo-transient-continuation (PTC) schemes for the numerical solution of semilinear elliptic partial differential equations, with possible singular perturbations. We will outline a residual reduction analysis within the framework of general Hilbert spaces, and, subsequently, employ the PTC-methodology in the context of finite element discretizations of semilinear boundary value problems. Our approach combines both a prediction-type PTC-method (for infinite dimensional problems) and an adaptive finite element discretization (based on a robust *a posteriori* residual analysis), thereby leading to a *fully adaptive PTC-Galerkin scheme*. Numerical experiments underline the robustness and reliability of the proposed approach for different examples.

## 1. INTRODUCTION AND PROBLEM FORMULATION

The focus of this paper is on the numerical approximation of semilinear elliptic partial differential equations (PDE), with possible singular perturbations. More precisely, for a fixed parameter  $\varepsilon > 0$  (possibly with  $\varepsilon \ll 1$ ), and a continuously differentiable function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , we consider the problem of finding a solution function  $u : \Omega \rightarrow \mathbb{R}$  which satisfies

$$-\varepsilon \Delta u = f(u) \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \quad (1)$$

Here,  $\Omega \subset \mathbb{R}^d$ , with  $d = 1$  or  $d = 2$ , is an open and bounded 1d interval or a 2d Lipschitz polygon, respectively. Problems of this type appear in a wide range of applications including, e.g., nonlinear reaction-diffusion in ecology and chemical models [8, 12, 14, 16, 17], economy [6], or classical and quantum physics [7, 21]. From an analysis point of view, semilinear elliptic boundary value problems (1) have been studied in detail by a number of authors over the last few decades; we refer, e.g., to the monographs [1, 18, 20] and the references therein. In particular, solutions of (1) are known to be typically not unique (even infinitely many solutions may exist), and, in the singularly perturbed case, to exhibit boundary layers, interior shocks, and (multiple) spikes. The existence of multiple solutions due to the nonlinearity of the problem and/or the appearance of singular effects are challenging issues when solving problems of this type numerically; see, e.g., [19, 23].

*Linearized Galerkin Methods.* There are, in general, two approaches when solving nonlinear differential equations numerically: Either the nonlinear PDE problem to be solved is first discretized; this leads to a nonlinear algebraic system. Or, alternatively, a local linearization procedure, resulting in a sequence of linear PDE problems, is applied; these linear problems are subsequently discretized by a suitable numerical approximation scheme. We emphasize that the latter approach enables the use of the large body of existing numerical analysis and computational techniques for *linear* problems (such as, e.g., the development of classical residual-based error bounds). The concept of approximating infinite dimensional nonlinear problems by appropriate *linear discretization schemes* has been studied by several authors in the recent past. For example, the approach presented in [10] (see also the work [9, 15]) combines fixed point linearization methods and Galerkin approximations in the context of strictly monotone problems. Similarly, in [2–4] (see also [13]),

---

MATHEMATICS INSTITUTE, UNIVERSITY OF BERN, CH-3012 SWITZERLAND

*E-mail addresses:* mario.amrein@hslu.ch, wihler@math.unibe.ch.

2010 *Mathematics Subject Classification.* 49M15, 58C15, 65N30.

*Key words and phrases.* Adaptive pseudo transient continuation method, dynamical system, steady states, semilinear elliptic problems, singularly perturbed problems, adaptive finite element methods.

the nonlinear PDE problems at hand are linearized by an (adaptive) Newton technique, and, subsequently, discretized by a linear finite element method. On a related note, the discretization of a sequence of linearized problems resulting from the local approximation of semilinear evolutionary problems has been investigated in [5]. In all of the works [2–5, 10], the key idea in obtaining fully adaptive discretization schemes is to provide a suitable interplay between the underlying linearization procedure and (adaptive) Galerkin methods; this is based on investing computational time into whichever of these two aspects is currently dominant.

*PTC-Approach.* In contrast to the classical Newton linearization method, the approach to be discussed in this work relies on a pseudo transient continuation procedure (see, e.g., [11, §6.4] for finite dimensional problems). The basis of this idea is to first interpret any solution  $u$  of the nonlinear equation  $F(u) = 0$ , where  $F$  is a given operator, as a steady state of the initial value problem

$$\dot{u} = F(u), \quad u(0) = u_0,$$

and, then, to discretize the dynamical system in time by means of the backward Euler method. Furthermore, the resulting sequence of nonlinear problems,  $u_{n+1} = u_n + t_n F(u_{n+1})$ ,  $n \geq 0$ , where  $t_n > 0$  is a given time step, is linearized with the aid of the Newton method. This scheme is termed PTC-method. On a local level, i.e., whenever the iteration is close enough to a solution point, the PTC-method turns into the standard Newton method. Otherwise, if the iteration is far away from a solution point, then the scheme can be interpreted as a continuation method. In a certain sense, the PTC-method can also be understood as an inexact Newton method. Following the methodology developed in the articles [3–5, 10], the present paper employs the idea of combining the PTC-linearization approach with adaptive  $\mathbb{P}_1$ -finite element methods (FEM). Our analysis will proceed along the lines of [11, §6.4], with the aim to provide an optimal residual reduction procedure in the local linearization process. Moreover, in order to address the issue of devising  $\varepsilon$ -robust *a posteriori* error estimates for the Galerkin discretizations, we employ the approach presented in [22].

*Outline.* The outline of this paper is as follows. In Section 2 we study the PTC-method within the context of general Hilbert spaces, and derive a residual reduction analysis. Subsequently, the purpose of Section 3 is the discretization of the resulting sequence of *linear* problems by the finite element method, and the development of an  $\varepsilon$ -robust *a posteriori* error analysis. The final estimate (Theorem 3.5) bounds the residual in terms of the (elementwise) finite element approximation (FEM-error) and the error caused by the linearization of the original problem. Then, in order to define a fully adaptive PTC-Galerkin scheme, we propose an interplay between the adaptive PTC-method and the adaptive finite element approach: More precisely, as the adaptive procedure is running, we either perform a PTC-step in accordance with the suggested prediction strategy (Section 2) or refine the current finite element mesh based on the *a posteriori residual* estimate (Section 3); this is carried out depending on which of the errors (FEM-error or PTC-error) is more dominant in the present iteration step. In Section 3.5 we provide a series of numerical experiments which show that the proposed scheme is reliable and  $\varepsilon$ -robust for reasonable choices of initial guesses. Finally, we add a few concluding remarks in Section 4.

*Problem Formulation.* In this paper, we suppose that a (not necessarily unique) solution  $u \in X := H_0^1(\Omega)$  of (1) exists; here, we denote by  $H_0^1(\Omega)$  the standard Sobolev space of functions in  $H^1(\Omega) = W^{1,2}(\Omega)$  with zero trace on  $\partial\Omega$ . Furthermore, signifying by  $X' = H^{-1}(\Omega)$  the dual space of  $X$ , and upon defining the map  $F_\varepsilon : X \rightarrow X'$  through

$$\langle F_\varepsilon(u), v \rangle := \int_{\Omega} \{f(u)v - \varepsilon \nabla u \cdot \nabla v\} \, d\mathbf{x} \quad \forall v \in X, \quad (2)$$

where  $\langle \cdot, \cdot \rangle$  signifies the dual pairing in  $X' \times X$ , the above problem (1) can be written as a nonlinear operator equation in  $X'$ :

$$F_\varepsilon(u) = 0,$$

for an unknown zero  $u \in X$ . For the purpose of defining the Newton linearization later on in this manuscript, we note that the Fréchet-derivative of  $F_\varepsilon$  at  $u \in X$  is given by

$$\langle F'_\varepsilon(u)w, v \rangle = \int_\Omega \{f'(u)wv - \varepsilon \nabla w \nabla v\} d\mathbf{x}, \quad v, w \in X,$$

where we write  $f'(u) = \partial_u f(u)$ . In addition, we introduce the inner product

$$(u, v)_X := \int_\Omega \{uv + \varepsilon \nabla u \cdot \nabla v\} d\mathbf{x}, \quad u, v \in V,$$

with induced norm on  $X$  given by

$$\|u\|_{\varepsilon, D} := \left( \varepsilon \|\nabla u\|_{0, D}^2 + \|u\|_{0, D}^2 \right)^{1/2}, \quad u \in H^1(D),$$

where  $\|\cdot\|_{0, D}$  denotes the  $L^2$ -norm on  $D$ . Frequently, for  $D = \Omega$ , the subindex ' $D$ ' will be omitted. Note that, in the case of  $f(u) = -u + g$ , with  $g \in L^2(\Omega)$ , i.e., when (1) is linear and strongly elliptic, the norm  $\|\cdot\|_{\varepsilon, \Omega}$  is a natural energy norm on  $X$ . As usual, for any  $\varphi \in X'$ , the dual norm is given by

$$\|\varphi\|_{X'} = \sup_{x \in X \setminus \{0\}} \frac{\langle \varphi, x \rangle}{\|x\|_\varepsilon}.$$

In what follows we shall use the abbreviation  $x \preccurlyeq y$  to mean  $x \leq cy$ , for a constant  $c > 0$  independent of the mesh size  $h$  and of  $\varepsilon > 0$ .

## 2. ABSTRACT FRAMEWORK IN HILBERT SPACES

In this section we briefly revisit a possible derivation of the PTC-scheme. Moreover, following along the lines of [11] we will discuss how residual reduction, based on a PTC-iteration-scheme, can be achieved within the context of general Hilbert spaces. To this end, let  $X$  be a real Hilbert space with inner product  $(\cdot, \cdot)_X$  and induced norm  $(x, x)_X^{1/2} = \|x\|_X$ . Furthermore, by  $\mathcal{L}(X; X')$ , we signify the space of all bounded linear operators from  $X$  into  $X'$ , with norm

$$\|L\|_{\mathcal{L}(X; X')} = \sup_{\substack{x \in X \\ \|x\|_X = 1}} \|L(x)\|_{X'},$$

for any  $L \in \mathcal{L}(X; X')$ .

**2.1. PTC-Scheme.** We take the view of dynamical systems, i.e., given a possibly nonlinear operator

$$F : X \rightarrow X',$$

we interpret any zero  $u_\infty \in X$  of  $F$ , i.e.,  $F(u_\infty) = 0$ , as a *steady state* of the dynamical system

$$u(0) = u_0, \quad (\dot{u}(t), v)_X = \langle F(u(t)), v \rangle \quad \forall v \in X, t > 0, \quad (3)$$

where we denote by  $\langle \cdot, \cdot \rangle$  the dual pairing in  $X' \times X$  as before, and  $u_0 \in X$  is a given initial guess. More precisely, we suppose that there exists a solution  $u : [0, \infty) \rightarrow X$  of (3) with  $\lim_{t \rightarrow \infty} u(t) = u_\infty$  in a suitable sense. Then, we discretize (3) in time using the backward Euler method, i.e.,

$$(u_{n+1}, v)_X = (u_n, v)_X + k_n \langle F(u_{n+1}), v \rangle \quad \forall v \in X, \quad n \geq 0, \quad (4)$$

where  $k_n > 0$  signifies the (possibly adaptively chosen) temporal step size. Introducing, for  $n \geq 0$ , an operator  $G_n : X \rightarrow X'$  by

$$\langle G_n(u), v \rangle := (u - u_n, v)_X - k_n \langle F(u), v \rangle \quad \forall v \in X,$$

we see that the zeros of  $G_n$  define the next update  $u_{n+1}$  in (4). Then, applying Newton's method to  $G_n$  yields a linear equation for an unknown increment  $\tilde{\delta}_n \in X$  such that

$$\langle G'_n(u_n) \tilde{\delta}_n, v \rangle = -\langle G_n(u_n), v \rangle \quad \forall v \in X,$$

and the update

$$u_{n+1} = u_n + \tilde{\delta}_n,$$

where  $G'_n$  denotes the Fréchet derivative of  $G_n$ . Equivalently, upon rescaling  $\delta_n = k_n^{-1} \tilde{\delta}_n$ , we have

$$(\delta_n, v)_X - k_n \langle F'(u_n) \delta_n, v \rangle = \langle F(u_n), v \rangle, \quad u_{n+1} = u_n + k_n \delta_n. \quad (5)$$

Incidentally, with  $\delta_n \rightharpoonup 0$  weakly in  $X$ , we obtain Newton's method as applied to  $F$ . In order to simplify notation we introduce, for given  $u \in X$  and  $t > 0$ , an additional operator

$$A[t; u] : X \rightarrow X',$$

which is defined by

$$x \in X : \quad \langle A[t; u]x, v \rangle := (x, v)_X - t \langle F'(u)x, v \rangle \quad \forall v \in X. \quad (6)$$

We can then rewrite (5) as

$$\langle A[k_n; u_n] \delta_n, v \rangle = \langle F(u_n), v \rangle \quad \forall v \in X, \quad u_{n+1} = u_n + k_n \delta_n. \quad (7)$$

For  $n = 0, 1, 2, \dots$ , with a given initial guess  $u_0 \in X$ , this iteration defines the PTC-scheme for the approximation of a zero of  $F$ . Evidently, in order to be able to solve for  $\delta_n$  in (7), the operator  $A[k_n; u_n]$  needs to be invertible.

**2.2. Residual Analysis.** The aim of this section is to derive a residual estimate which paves the way for a residual reduction time stepping strategy. It is based on the following structural assumptions on the derivative of  $F$ :

(a) For given  $u_0 \in X$ , there exists a constant  $\mu = \mu(u_0) > 0$  such that

$$\sup_{\substack{x \in X \\ \|x\|_X = 1}} \langle F'(u_0)x, x \rangle \leq -\mu. \quad (A.1)$$

(b) There is a constant  $L \geq 0$  such that there holds the Lipschitz property

$$\|F'(x) - F'(y)\|_{\mathcal{L}(X; X')} \leq L \|x - y\|_X \quad \forall x, y \in X. \quad (A.2)$$

**Proposition 2.1.** *Let  $u_0 \in X$  such that  $F'(u_0) \in \mathcal{L}(X; X')$ . If (A.1) holds, and if  $F(u_0) \in X'$ , then the linear problem*

$$A[t; u_0](u(t) - u_0) = tF(u_0) \quad (8)$$

*has a unique solution  $u(t) \in X$  for any  $t > 0$ .*

*Proof.* We apply the Lax-Milgram Lemma. In particular, we show that  $A[t; u_0]$  is coercive and bounded on  $X$ . Indeed, for all  $v \in X$ , we have

$$\langle A[t; u_0]v, v \rangle = (v, v)_X - t \langle F'(u_0)v, v \rangle \geq (1 + \mu t) \|v\|_X^2, \quad (9)$$

which proves coercivity. Moreover, for  $v, w \in X$ , we have

$$|\langle A[t; u_0]v, w \rangle| \leq \|v\|_X \|w\|_X + t \|F'(u_0)v\|_{X'} \|w\|_X.$$

Since  $F'(u_0)$  is bounded, we deduce the boundedness of  $A[t; u_0]$ . This completes the proof.  $\square$

In order to devise a residual reduction analysis, we insert two preparatory results.

**Lemma 2.2.** *If (A.1) is satisfied, then we have*

$$\|A[t; u_0]^{-1}\|_{\mathcal{L}(X'; X)} \leq \frac{1}{1 + t\mu}. \quad (10)$$

*Moreover, if  $F(u_0) \in X'$ , the estimate*

$$\|u(t) - u_0\|_X \leq \frac{t}{1 + t\mu} \|F(u_0)\|_{X'} \quad (11)$$

*holds true, where  $u(t)$ ,  $t \geq 0$ , is the solution from (8).*

*Proof.* From (9), we readily arrive at

$$\|A[t; u_0]v\|_{X'} \geq \|v\|_X (1 + t\mu) \quad \forall v \in X,$$

from which we deduce (10). Furthermore, the second bound results by definition of  $A[t; u_0]$  in (6) with  $v = u(t) - u_0$ , and from (A.1). Indeed,

$$\begin{aligned} \|u(t) - u_0\|_X^2 &= \langle A[t; u_0](u(t) - u_0), u(t) - u_0 \rangle + t \langle F'(u_0)(u(t) - u_0), u(t) - u_0 \rangle \\ &= t \langle F(u_0), u(t) - u_0 \rangle + t \langle F'(u_0)(u(t) - u_0), u(t) - u_0 \rangle \\ &\leq t \|F(u_0)\|_{X'} \|u(t) - u_0\|_X - t\mu \|u(t) - u_0\|_X^2, \end{aligned} \quad (12)$$

which immediately implies (11).  $\square$

**Lemma 2.3.** *If  $u$  from (8) is differentiable for any  $t \geq 0$ , then it holds that*

$$A[t; u_0]\dot{u}(t) = F(u_0) + F'(u_0)(u(t) - u_0), \quad (13)$$

as well as

$$tA[t; u_0]\dot{u}(t) = (u(t) - u_0, \cdot)_X \quad (14)$$

in  $X'$ .

*Proof.* Recalling (6), we observe that

$$\frac{d}{dt} A[t; u_0]v = -F'(u_0)v \quad \forall v \in X,$$

in  $X'$ . Then, differentiating (8) with respect to  $t$  implies

$$A[t; u_0]\dot{u}(t) - F'(u_0)(u(t) - u_0) = F(u_0),$$

which yields (13). Furthermore, multiplying this equality by  $t$ , and applying the definition of  $A[t; u_0]$  from (6), it follows that

$$tA[t; u_0]\dot{u}(t) = tF(u_0) + (u(t) - u_0, \cdot)_X - A[t; u_0](u(t) - u_0)$$

in  $X'$ . Using (8) gives (14).  $\square$

Following along the lines of [11] there holds the ensuing residual reduction result.

**Theorem 2.4.** *Under the assumptions in Lemmas 2.2 and 2.3, and if (A.2) holds, then we have*

$$\|F(u(t))\|_{X'} \leq \gamma(t) \|F(u_0)\|_{X'}, \quad (15)$$

with

$$\gamma(t) := \frac{1}{1 + t\mu} \left( 1 + \frac{Lt^2}{2(1 + t\mu)} \|F(u_0)\|_{X'} \right) > 0,$$

for any  $t \geq 0$ .

*Proof.* For  $t \geq 0$ , there holds

$$\begin{aligned} F(u(t)) &= F(u_0) + \int_0^t F'(u(s))\dot{u}(s) \, ds \\ &= F(u_0) + F'(u_0)(u(t) - u_0) + \int_0^t (F'(u(s)) - F'(u_0))\dot{u}(s) \, ds. \end{aligned}$$

Involving (13), we infer that

$$F(u(t)) = A[t; u_0]\dot{u}(t) + \int_0^t (F'(u(s)) - F'(u_0))\dot{u}(s) \, ds.$$

Therefore, we have

$$\|F(u(t))\|_{X'} \leq \|A[t; u_0]\dot{u}(t)\|_{X'} + \int_0^t \|(F'(u(s)) - F'(u_0))\dot{u}(s)\|_{X'} \, ds.$$

Employing (A.2) and applying (14), we arrive at

$$t \|F(u(t))\|_{X'} \leq \|u(t) - u_0\|_X + Lt \int_0^t \|u(s) - u_0\|_X \|\dot{u}(s)\|_X \, ds. \quad (16)$$

Moreover, again from (14), we notice that

$$s \|\dot{u}(s)\|_X \leq \|A[s; u_0]^{-1}\|_{\mathcal{L}(X'; X)} \|u(s) - u_0\|_X, \quad s \geq 0, \quad (17)$$

and, hence, by virtue of Lemma 2.2, we obtain

$$s \|u(s) - u_0\|_X \|\dot{u}(s)\|_X \leq \frac{1}{1 + s\mu} \|u(s) - u_0\|_X^2 \leq \frac{s^2}{(1 + s\mu)^3} \|F(u_0)\|_{X'}^2.$$

Combining this with (16), and using Lemma 2.2 once more, leads to

$$\begin{aligned} t \|F(u(t))\|_{X'} &\leq \frac{t}{1 + t\mu} \|F(u_0)\|_{X'} + Lt \|F(u_0)\|_{X'}^2 \int_0^t \frac{s}{(1 + s\mu)^3} ds \\ &= \frac{t}{1 + t\mu} \|F(u_0)\|_{X'} + \frac{Lt^3}{2(1 + \mu t)^2} \|F(u_0)\|_{X'}^2. \end{aligned}$$

This completes the proof.  $\square$

**Remark 2.5.** Referring to (12), we see that

$$\|u(t) - u_0\|_X^2 \leq t \langle F(u_0), u(t) - u_0 \rangle - t\mu \|u(t) - u_0\|_X^2.$$

Hence, whenever there holds  $t \langle F(u_0), u(t) - u_0 \rangle \leq \|u(t) - u_0\|_X^2$ , it follows that  $\mu \leq 0$  (as long as there is  $t > 0$  with  $u(t) \neq u_0$ ). In particular, assumption (A.1) is not fulfilled in this case. We may therefore assume that

$$\|u(t) - u_0\|_X^2 < t \langle F(u_0), u(t) - u_0 \rangle,$$

for  $t > 0$ , and  $u(t) \neq u_0$ .

From (15) it follows that the residual decreases, i.e.,  $\|F(u(t))\|_{X'} < \|F(u_0)\|_{X'}$ , if  $\gamma(t) \in (0, 1)$ . For  $t > 0$ , this happens if there holds

$$\left( \frac{L}{2} \|F(u_0)\|_{X'} - \mu^2 \right) t < \mu \quad (t > 0). \quad (18)$$

Therefore, if

$$\frac{L}{2} \|F(u_0)\|_{X'} \leq \mu^2,$$

then any value of  $t > 0$  will lead to a reduction of the residual. Otherwise, (18) can be satisfied as long as  $t$  is chosen sufficiently small; in the special case that  $L \|F(u_0)\|_{X'} > \mu^2$ , it is elementary to verify that  $\gamma(t)$  attains its minimum for

$$t^* = \frac{\mu}{L \|F(u_0)\|_{X'} - \mu^2}.$$

**2.3. Pseudo Time Stepping.** In terms of the PTC-scheme (7), for  $n \geq 0$ , our previous discussion translates into

$$\|F(u_{n+1})\|_{X'} \leq \gamma_n \|F(u_n)\|_{X'},$$

with a reduction constant

$$\gamma_n = \frac{1}{1 + k_n \mu} \left( 1 + \frac{L k_n^2}{2(1 + k_n \mu)} \|F(u_n)\|_{X'} \right) > 0;$$

cf. Theorem 2.4. If

$$\frac{L}{2} \|F(u_n)\|_{X'} \leq \mu^2,$$

then any choice of  $k_n > 0$  will imply that  $\gamma_n \in (0, 1)$ . Otherwise, for  $k_n$  sufficiently small so that

$$\left( \frac{L}{2} \|F(u_n)\|_{X'} - \mu^2 \right) k_n < \mu,$$

it holds that  $\gamma_n < 1$ . In particular, if  $L \|F(u_n)\|_{X'} > \mu^2$ , then

$$k_n^* = \frac{\mu}{L \|F(u_n)\|_{X'} - \mu^2} \quad (19)$$

results in a minimal value of  $\gamma_n$ . For this value of  $k_n$ , we apply (11) to infer the bound

$$\|u_{n+1} - u_n\|_X \leq \frac{k_n^*}{1 + k_n^* \mu} \|F(u_n)\|_{X'} = \frac{\mu}{L}.$$

Letting  $\delta_n = (u_{n+1} - u_n)/k_n^*$  be the increment in the PTC-iteration (5), this leads to  $\|\delta_n\|_X \leq \mu/(k_n^* L)$ , and, therefore,

$$k_n^* \leq \frac{\mu}{L \|\delta_n\|_X}. \quad (20)$$

This upper bound does not contain any dual norms, and can, thus, be employed as an approximation of  $k_n^*$  in practice.

**Remark 2.6.** In an effort to replace (20) by a computationally even more feasible bound (not involving the possibly unspecified constants  $\mu$  and  $L$ ), we proceed again along the lines of [11]. As in (12), for  $k_n > 0$ , we have

$$\|\delta_n\|_X^2 \leq \langle F(u_n), \delta_n \rangle - \mu k_n \|\delta_n\|_X^2.$$

This motivates to define the computable quantity

$$\mu_n := \frac{\langle F(u_n), \delta_n \rangle - \|\delta_n\|_X^2}{k_n \|\delta_n\|_X^2} \geq \mu > 0.$$

Furthermore, similarly as in the proof of Theorem 2.4, we note that

$$\langle F(u_{n+1}), \delta_n \rangle = \|\delta_n\|_X^2 + \int_{t_n}^{t_{n+1}} \langle (F'(u(s)) - F'(u_n))\dot{u}(s), \delta_n \rangle ds,$$

where, for  $i \geq 1$ , we let  $t_i = \sum_{j=0}^{i-1} k_j$ . Then, by means of (A.2), it follows that

$$\begin{aligned} \frac{|\langle F(u_{n+1}), \delta_n \rangle - \|\delta_n\|_X^2|}{\|\delta_n\|_X} &\leq \int_{t_n}^{t_{n+1}} \|(F'(u(s)) - F'(u_n))\dot{u}(s)\|_{X'} ds \\ &\leq L \int_{t_n}^{t_{n+1}} \|u(s) - u_n\|_X \|\dot{u}(s)\|_X ds. \end{aligned}$$

Furthermore, using (17) and employing Lemma 2.2, this transforms into

$$\frac{|\langle F(u_{n+1}), \delta_n \rangle - \|\delta_n\|_X^2|}{\|\delta_n\|_X} \leq L \int_{t_n}^{t_{n+1}} \frac{(s - t_n)^{-1}}{1 + (s - t_n)\mu} \|u(s) - u_n\|_X^2 ds.$$

Approximating the integral with the aid of the trapezoidal rule, and recalling that  $u_{n+1} - u_n = k_n \delta_n$ , cf. (5), yields

$$\frac{|\langle F(u_{n+1}), \delta_n \rangle - \|\delta_n\|_X^2|}{\|\delta_n\|_X} \leq \frac{L}{2} k_n^2 \|\delta_n\|_X^2 + \mathcal{O}(k_n^4).$$

We then define

$$\mathbf{L}_n := \frac{2|\langle F(u_{n+1}), \delta_n \rangle - \|\delta_n\|_X^2|}{k_n^2 \|\delta_n\|_X^3} \leq L + \mathcal{O}(k_n^2).$$

Replacing  $\mu$  and  $L$  in (20) by  $\mu_n$  and  $\mathbf{L}_n$ , respectively, we are led to introduce the following pseudo time step

$$\mathbf{k}_n^* = \frac{k_n}{2} \cdot \left| \frac{\langle F(u_n), \delta_n \rangle - \|\delta_n\|_X^2}{\langle F(u_{n+1}), \delta_n \rangle - \|\delta_n\|_X^2} \right|, \quad (21)$$

which does not require explicit knowledge on  $\mu$  and  $L$ .

### 3. APPLICATION TO SEMILINEAR PROBLEMS

In this section, we will apply the abstract setting from the previous section to the semilinear problem (1), with  $F = F_\varepsilon$  from (2).

**3.1. PTC-Linearization.** For  $u_n \in X$  and  $k_n > 0$ , the PTC-method (5) is to find  $\delta_n \in X$  such that

$$a_\varepsilon(u_n, k_n; \delta_n, v) = \ell_\varepsilon(u_n; v) \quad \forall v \in X, \quad (22)$$

and  $u_{n+1} = u_n + k_n \delta_n$ , where, for fixed  $u \in X$ ,  $t > 0$ , we consider the bilinear form

$$a_\varepsilon(u, t; \delta, v) := (\delta, v)_X - t \int_\Omega \{f'(u)\delta v - \varepsilon \nabla \delta \cdot \nabla v\} \, d\mathbf{x}, \quad \delta, v \in X,$$

as well as the linear form

$$\ell_\varepsilon(u; v) := \int_\Omega \{f(u)v - \varepsilon \nabla u \cdot \nabla v\} \, d\mathbf{x}, \quad v \in X.$$

Throughout, for given  $u_n$ ,  $n \geq 0$ , we assume that (5) has a unique solution  $u_{n+1}$ . In fact, this property can be made rigorous if certain assumptions on the nonlinearity  $f$  are satisfied. This will be addressed in the ensuing two propositions.

**Proposition 3.1.** *If  $\sigma_f := \sup_{x \in \mathbb{R}} f'(x) < \varepsilon C_{\text{Poinc}}^{-2}$ , where  $C_{\text{Poinc}} = C_{\text{Poinc}}(\Omega)$  is the constant in the Poincaré inequality on  $\Omega$ ,*

$$\|w\|_0 \leq C_{\text{Poinc}} \|\nabla w\|_0, \quad \forall w \in X, \quad (23)$$

then (A.1) is satisfied with

$$\mu = \frac{\varepsilon C_{\text{Poinc}}^{-2} - \sigma_f}{\varepsilon C_{\text{Poinc}}^{-2} + 1} > 0.$$

*Proof.* Let us set

$$\zeta := \frac{1 + \sigma_f}{\varepsilon C_{\text{Poinc}}^{-2} + 1}.$$

By our assumptions, there holds that  $\zeta < 1$ . Then, with (23), for  $u, v \in X$ , we have

$$\begin{aligned} \langle F'_\varepsilon(u)v, v \rangle &= (\zeta - 1)\varepsilon \|\nabla v\|_0^2 - \zeta \varepsilon \|\nabla v\|_0^2 + \int_\Omega f'(u)v^2 \, d\mathbf{x} \\ &\leq (\zeta - 1)\varepsilon \|\nabla v\|_0^2 + \int_\Omega \{f'(u) - \zeta C_{\text{Poinc}}^{-2}\varepsilon\}v^2 \, d\mathbf{x} \\ &\leq (\zeta - 1) \|v\|_\varepsilon^2 = -\mu \|v\|_\varepsilon^2. \end{aligned}$$

Hence, (A.1) is verified.  $\square$

**Remark 3.2.** Within a given PTC-iteration, for  $n \geq 0$ , the proof of the above result reveals that  $\sigma_f$  can be replaced by the possibly sharper value  $\sigma_f := \sup_\Omega f'(u_n)$ .

**Proposition 3.3.** *If  $f'$  is globally Lipschitz continuous with Lipschitz constant  $L_{f'}$ , that is,*

$$|f'(u_1) - f'(u_2)| \leq L_{f'} |u_1 - u_2| \quad \forall u_1, u_2 \in \mathbb{R}, \quad (24)$$

then (A.2) is fulfilled with  $L := CL_{f'}\varepsilon^{-1}$ , where  $C > 0$  is a constant only depending on  $\Omega$ .

*Proof.* For  $u_1, u_2, w, v \in X$  there holds

$$|\langle (F'(u_1) - F'(u_2))w, v \rangle| \leq \|(f'(u_1) - f'(u_2))w\|_{L^1(\Omega)} \|v\|_0.$$

Employing [4, Lemma A.1], and applying the Lipschitz continuity (24), we obtain

$$\begin{aligned} |\langle (F'(u_1) - F'(u_2))w, v \rangle| &\leq C \|f'(u_1) - f'(u_2)\|_0 \|\nabla w\|_0 \|\nabla v\|_0 \\ &\leq CL_{f'} \|u_1 - u_2\|_0 \|\nabla w\|_0 \|\nabla v\|_0 \\ &\leq CL_{f'} \varepsilon^{-1} \|u_1 - u_2\|_\varepsilon \|w\|_\varepsilon \|v\|_\varepsilon, \end{aligned}$$

for a constant  $C > 0$  only depending on  $\Omega$ . This verifies (A.2).  $\square$

**Remark 3.4.** If the assumptions in the above Propositions 3.1 and 3.3 are satisfied, and if for given  $u_n \in X$  there holds  $f(u_n) \in L^2(\Omega)$ , then the linear problem (22) has a unique solution  $\delta_n \in X$ ; cf. Proposition 2.1.



**3.2. PTC-Galerkin Discretization.** In order to provide a numerical approximation of (1), we will discretize the *linear* weak formulation (22) by means of a finite element method, which, in combination with the PTC-iteration, constitutes a PTC-Galerkin approximation scheme. Furthermore, we shall derive *a posteriori* residual estimates for the finite element discretization which allow for an adaptive refinement of the meshes in each PTC-step. This, together with the adaptive prediction strategy from Section 2, leads to a fully adaptive PTC-Galerkin discretization method for (1).

**3.2.1. Finite Element Meshes and Spaces.** Let  $\mathcal{T}^h = \{T\}_{T \in \mathcal{T}^h}$  be a regular and shape-regular mesh partition of  $\Omega$  into disjoint open simplices, i.e., any  $T \in \mathcal{T}^h$  is an affine image of the (open) reference simplex  $\hat{T} = \{\hat{x} \in \mathbb{R}_+^d : \sum_{i=1}^d \hat{x}_i < 1\}$ . By  $h_T = \text{diam}(T)$  we signify the element diameter of  $T \in \mathcal{T}^h$ , and by  $h = \max_{T \in \mathcal{T}^h} h_T$  the mesh size. Furthermore, by  $\mathcal{E}^h$  we denote the set of all interior mesh nodes for  $d = 1$  and interior (open) edges for  $d = 2$  in  $\mathcal{T}^h$ . In addition, for  $T \in \mathcal{T}^h$ , we let  $\mathcal{E}^h(T) = \{E \in \mathcal{E}^h : E \subset \partial T\}$ . For  $E \in \mathcal{E}^h$ , we let  $h_E$  be the mean of the lengths of the adjacent elements in 1d, and the length of  $E$  in 2d. Let us also define the following two quantities:

$$\alpha_T := \min(1, \varepsilon^{-1/2} h_T), \quad \alpha_E := \min(1, \varepsilon^{-1/2} h_E), \quad (25)$$

for  $T \in \mathcal{T}^h$  and  $E \in \mathcal{E}^h$ , respectively.

We consider the finite element space of continuous, piecewise linear functions on  $\mathcal{T}^h$  with zero trace on  $\partial\Omega$ , given by

$$V_0^h := \{\varphi \in H_0^1(\Omega) : \varphi|_T \in \mathbb{P}_1(T) \forall T \in \mathcal{T}^h\},$$

respectively, where  $\mathbb{P}_1(T)$  is the standard space of all linear polynomial functions on  $T$ .

**3.2.2. Linear Finite Element Discretization.** For given  $k_n > 0$  and  $u_n^h \in V_0^h$ ,  $n \geq 0$ , we consider the finite element approximation of (22), which is to find  $\delta_n^h \in V_0^h$  such that

$$a_\varepsilon(u_n^h, k_n; \delta_n^h, v) = \ell_\varepsilon(u_n^h; v) \quad \forall v \in V_0^h; \quad (26)$$

for  $n = 0$ , the function  $u_0^h \in V_0^h$  is a prescribed initial guess. Introducing the linearization operator

$$\mathbb{T}_f(u) := f(u_n^h) + f'(u_n^h)(u - u_n^h),$$

as well as

$$u_{n+1}^h := u_n^h + k_n \delta_n^h,$$

and rearranging terms, (26) can be rewritten as

$$\int_{\Omega} \varepsilon \nabla u_{n+1}^h \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} (\mathbb{T}_f(u_{n+1}^h) - \delta_n^h) v \, d\mathbf{x}, \quad (27)$$

for any  $v \in V_0^h$ .

**3.3. A Posteriori Residual Analysis.** The aim of this section is to derive *a posteriori* residual bounds for the linearized FEM (26).

**3.3.1. A Posteriori Residual Bound.** In order to measure the discrepancy between the finite element discretization (26) and the original problem (1), a natural quantity to bound is the residual  $F_\varepsilon(u_{n+1}^h)$  in  $X'$ . Let  $\mathcal{I}^h : H_0^1(\Omega) \rightarrow V_0^h$  be the quasi-interpolation operator of Clément (see, e.g., [4, Corollary 4.2]). Then, testing (27) with  $\mathcal{I}^h v \in V_0^h$ , for an arbitrary  $v \in X$ , implies that

$$\int_{\Omega} \varepsilon \nabla u_{n+1}^h \cdot \nabla \mathcal{I}^h v \, d\mathbf{x} = \int_{\Omega} (\mathbb{T}_f(u_{n+1}^h) - \delta_n^h) \mathcal{I}^h v \, d\mathbf{x}.$$

Then, there holds the identity

$$\begin{aligned} \langle F_\varepsilon(u_{n+1}^h), v \rangle &= \int_{\Omega} \varepsilon \nabla u_{n+1}^h \cdot \nabla (\mathcal{I}^h v - v) \, d\mathbf{x} + \int_{\Omega} (\mathbb{T}_f(u_{n+1}^h) - \delta_n^h) (v - \mathcal{I}^h v) \, d\mathbf{x} \\ &\quad - \int_{\Omega} \{ \mathbb{T}_f(u_{n+1}^h) - \delta_n^h - f(u_{n+1}^h) \} v \, d\mathbf{x}, \end{aligned}$$

for any  $v \in X$ . Integrating by parts in the first term on the right-hand side, recalling the fact that  $(v - \mathcal{I}^h v) = 0$  on  $\partial\Omega$ , and applying some elementary calculations, yields that

$$\langle \mathbf{F}_\varepsilon(u_{n+1}^h), v \rangle = \sum_{E \in \mathcal{E}^h} a_E + \sum_{T \in \mathcal{T}^h} (b_T - c_T),$$

where

$$\begin{aligned} a_E &:= \int_E \varepsilon \llbracket \nabla u_{n+1}^h \rrbracket (\mathcal{I}^h v - v) \, ds, & c_T &:= \int_T \{ \mathbb{T}_f(u_{n+1}^h) - \delta_n^h - f(u_{n+1}^h) \} v \, d\mathbf{x}, \\ b_T &:= \int_T \{ \varepsilon \Delta u_{n+1}^h + \mathbb{T}_f(u_{n+1}^h) - \delta_n^h \} (v - \mathcal{I}^h v) \, d\mathbf{x}, \end{aligned}$$

with  $E \in \mathcal{E}^h$ ,  $T \in \mathcal{T}^h$ . Here, for any edge  $E = \partial T^\sharp \cap \partial T^\flat \in \mathcal{E}^h$  shared by two neighboring elements  $T^\sharp, T^\flat \in \mathcal{T}^h$ , where  $\mathbf{n}^\sharp$  and  $\mathbf{n}^\flat$  signify the unit outward vectors on  $\partial T^\sharp$  and  $\partial T^\flat$ , respectively, we denote by

$$\llbracket \nabla u_{n+1}^h \rrbracket(\mathbf{x}) = \lim_{t \rightarrow 0^+} \nabla u_{n+1}^h(\mathbf{x} + t\mathbf{n}^\sharp) \cdot \mathbf{n}^\sharp + \lim_{t \rightarrow 0^+} \nabla u_{n+1}^h(\mathbf{x} + t\mathbf{n}^\flat) \cdot \mathbf{n}^\flat, \quad \mathbf{x} \in E,$$

the jump across  $E$ . Then, for  $T \in \mathcal{T}^h$ , defining the linearization residual

$$R_{n,T} := \|\mathbb{T}_f(u_{n+1}^h) - \delta_n^h - f(u_{n+1}^h)\|_{0,T}, \quad (28)$$

as well as the FEM approximation residual

$$\eta_{n,T}^2 := \alpha_T^2 \|\varepsilon \Delta u_{n+1}^h + \mathbb{T}_f(u_{n+1}^h) - \delta_n^h\|_{0,T}^2 + \frac{1}{2} \sum_{E \in \mathcal{E}^h(T)} \varepsilon^{-1/2} \alpha_E \|\varepsilon \llbracket \nabla u_{n+1}^h \rrbracket\|_{0,E}^2, \quad (29)$$

with  $\alpha_T$  and  $\alpha_E$  from (25), we proceed along the lines of the proof of [4, Theorem 4.4] in order to obtain the following result.

**Theorem 3.5.** *For  $n \geq 0$  there holds the upper a posteriori residual bound*

$$\|\mathbf{F}(u_{n+1}^h)\|_{X'}^2 \leq R_{n,\Omega}^2 + \sum_{T \in \mathcal{T}^h} \eta_{n,T}^2, \quad (30)$$

with  $R_{n,\Omega}$  and  $\eta_{n,T}$ ,  $T \in \mathcal{T}^h$ , from (28) and (29), respectively.

**Remark 3.6.** Following our approach in [4, Theorem 4.5], under certain conditions on the non-linearity  $f$ , it can be shown that the right-hand side of the above bound (30) is equivalent to the error norm  $\|u - u_{n+1}^h\|_{\varepsilon,\Omega}$ .

**Remark 3.7.** In addition to the upper estimate in the above Theorem 3.5, we notice that local lower *a posteriori* residual bounds can be established for the proposed PTC-Galerkin method as well. Indeed, this can be accomplished similarly to our analysis in [4, §4.4.2] (see also [22]), which is based on the application of standard bubble function techniques.

**3.4. A Fully Adaptive PTC-Galerkin Algorithm.** We will now propose a procedure that will combine the PTC-method presented in Section 2 with an automatic finite element mesh refinement strategy. More precisely, based on the *a posteriori* residual bound from Theorem 3.5, the main idea of our approach is to provide an interplay between PTC-iterations and adaptive mesh refinements which is based on monitoring the two residuals in (28) and (29), and on acting according to whatever quantity is dominant in the current computations. We make the assumption that the PTC-Galerkin sequence  $\{u_{n+1}^h\}_{n \geq 0}$  given by (26), with step size  $k_n^*$  from (19) (or  $k_n^*$  from (21)) is well-defined as long as the iterations are being performed. The individual computational steps are summarized in Algorithm 1.

**Algorithm 1** Fully-adaptive PTC-Galerkin method

---

```

1: Given a parameter  $\theta > 0$ , a (coarse starting) triangulation  $\mathcal{T}^h$  of  $\Omega$ , an initial step size  $k_0 > 0$ ,
   a maximal number of degrees of freedom  $\text{DOF}_{\max}$ , and an initial guess  $u_0^h \in V_0^h$ . Set  $n \leftarrow 0$ .
2: while  $\text{DOF} \leq \text{DOF}_{\max}$  do
3:   Compute the FEM solution  $u_{n+1}^h$  from (26) on the mesh  $\mathcal{T}^h$ .
4:   Evaluate the corresponding residual indicators  $\eta_{n,T}$ ,  $T \in \mathcal{T}^h$ , and  $R_{n,\Omega}$  from (28) and (29),
      respectively.
5:   if

$$R_{n,\Omega}^2 \leq \theta \sum_{T \in \mathcal{T}^h} \eta_{n,T}^2$$

      then
6:     refine the mesh  $T \in \mathcal{T}^h$  adaptively based on the elementwise residual indicators  $\eta_{n,T}$ ,
         $T \in \mathcal{T}^h$  from Theorem 3.5, and go back to step (2:) with the previously computed
        solution  $u_{n+1}^h$  as interpolated on the refined mesh;
7:   else
8:     perform another PTC-step based on the new step size  $k_n = k_n^*$  as proposed in (21) and
        go back to (3:).
9:   end if
10:  set  $n \leftarrow n + 1$ .
11: end while

```

---

**3.5. Numerical Experiments.** We will now illustrate and test the above fully adaptive Algorithm 1 with two numerical experiments in 2d. The linear systems resulting from the finite element discretization (27) are solved by means of a direct solver.

**Example 3.8.** Let us consider first the Sine-Gordon type problem

$$-\varepsilon \Delta u = -\sin(u) - u + 1, \text{ in } \Omega = (0, 1)^2, \quad u = 0 \text{ on } \partial\Omega.$$

Here,  $f(u) = -\sin(u) - u + 1$ , and  $f'(u) = -\cos(u) - 1$ . In particular, by application of Proposition 3.1, we observe that the structural assumptions (A.1) and (A.2) are fulfilled. Neglecting the boundary conditions for a moment, one observes that the unique positive zero  $u \approx 0.51$  of  $f(u)$  is a solution of the PDE. We therefore expect boundary layers along  $\partial\Omega$ ; see Figure 1 (right). Moreover, the focus of this experiment is on the robustness of the *a posteriori* residual bound (30) with respect to the singular perturbation parameter  $\varepsilon$  as  $\varepsilon \rightarrow 0$ . Starting from the initial mesh depicted in Figure 1 (left) with  $u_0^h(1/2, 1/2) \approx 1/2$ , we test the fully adaptive PTC-Galerkin Algorithm 1 for different choices of  $\varepsilon = \{10^{-i}\}_{i=0}^9$ . In Algorithm 1 the parameters are chosen to be  $\theta = 0.5$  and  $k_0 = 1$ . As  $\varepsilon \rightarrow 0$  the resulting solutions feature ever stronger boundary layers; see Figure 1 (right). The performance data in Figure 2 (left) shows that the residuals decay, firstly, robust in  $\varepsilon$ , and, secondly, of (optimal) order  $1/2$  with respect to the number of degrees of freedom.

**Example 3.9.** Finally, we turn to the well-known nonlinear Ginzburg-Landau equation on the square  $\Omega = (-1, 1)^2$  given by

$$-\varepsilon \Delta u = u(1 - u^2) \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega.$$

Clearly  $u \equiv 0$  is a solution. In addition, any solution  $u$  appears pairwise as  $-u$  is obviously a solution also. Again, neglecting the boundary conditions for a moment, we observe that  $u \equiv 1$  and  $u \equiv -1$  are solutions of the PDE. We therefore expect boundary layers along  $\partial\Omega$ , and possibly within the domain  $\Omega$ ; see Figure 3 (right). Here we always start from the initial mesh depicted in Figure 3 (left) with  $u_0^h \equiv 1$  on the interior nodes. Again we test the fully adaptive PTC-Galerkin Algorithm 1 for different choices of  $\varepsilon = \{10^{-i}\}_{i=0}^9$ . The parameters are still chosen to be  $\theta = 0.5$  and  $k_0 = 1$ . As in Example 3.8, for  $\varepsilon \rightarrow 0$  the resulting solution feature ever stronger boundary layers; see Figure 3 (right). In addition, from the performance data given in Figure 2 (right) we observe that the residuals decay again robust in  $\varepsilon$ . Finally we notice convergence of (optimal) order  $1/2$  with respect to the number of degrees of freedom. We remark that, although (A.1)

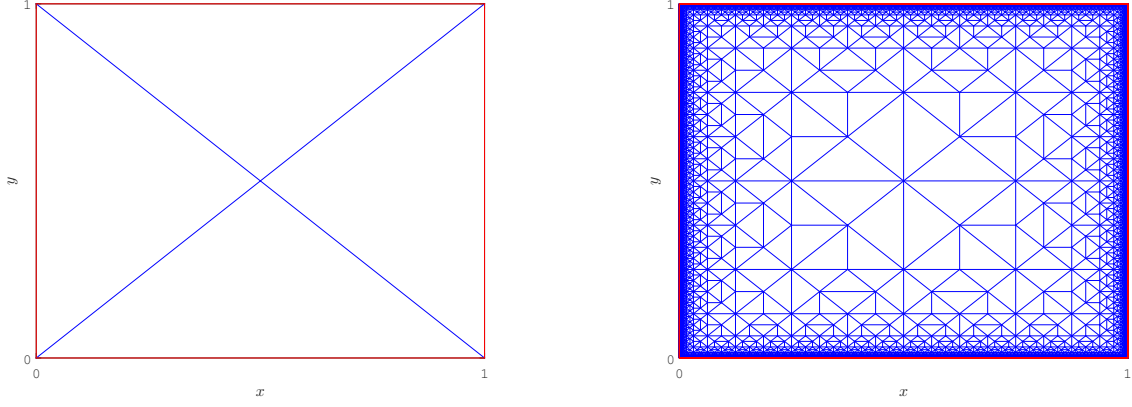


FIGURE 1. Example 3.8 for  $\varepsilon = 10^{-7}$ : Initial mesh (left), and the adaptively refined mesh resolving the solution (right).

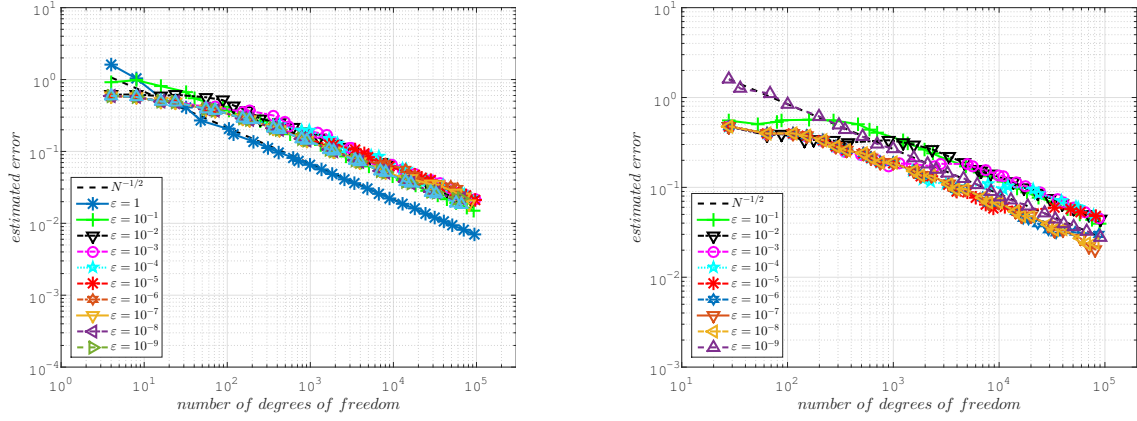


FIGURE 2. Estimated residuals for different choices of  $\varepsilon$ . On the left for Example 3.8 and on the right for Example 3.9.

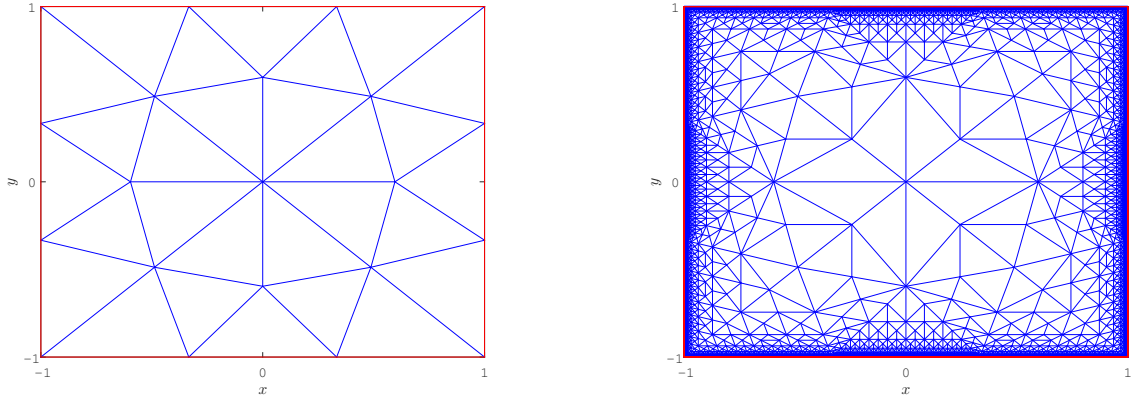


FIGURE 3. Example 3.9 for  $\varepsilon = 10^{-7}$ : Initial mesh (left), and the adaptively refined mesh resolving the solution (right).

and (A.2) are not necessarily satisfied for this problem, our fully adaptive PTC-Galerkin approach still delivers good results.

## 4. CONCLUSIONS

The aim of this paper was to introduce a reliable and computationally feasible procedure for the numerical solution of semilinear elliptic boundary value problems, with possible singular perturbations. The key idea is to combine adaptive step size control for the PTC-method with an automatic mesh refinement finite element procedure. Furthermore, the sequence of linear problems resulting from the application of pseudo transient continuation and Galerkin discretization is treated by means of a robust (with respect to the singular perturbations) *a posteriori* residual analysis and a corresponding adaptive mesh refinement process. Our numerical experiments clearly illustrate the ability of our approach to reliably find solutions reasonably close to the initial guesses, and to robustly resolve the singular perturbations at an optimal rate.

## REFERENCES

1. A. Ambrosetti and A. Malchiodi, *Perturbation methods and semilinear elliptic problems on  $\mathbf{R}^n$* , Progress in Mathematics, vol. 240, Birkhäuser Verlag, Basel, 2006.
2. M. Amrein, J.M. Melenk, and T. P. Wihler, *An hp-Adaptive Newton–Galerkin Finite Element Procedure for Semilinear Methods for Semilinear Boundary Value Problems*, Tech. report, <http://arxiv.org>, 2016.
3. M. Amrein and T. P. Wihler, *An adaptive Newton-method based on a dynamical systems approach*, Communications in Nonlinear Science and Numerical Simulation **19** (2014), no. 9, 2958–2973.
4. ———, *Fully Adaptive Newton–Galerkin Methods for Semilinear Elliptic Partial Differential Equations*, SIAM J. Sci. Comput. **37** (2015), no. 4, A1637–A1657.
5. ———, *Fully adaptive Newton–Galerkin time stepping methods for singularly perturbed parabolic evolution equations*, Tech. report, <http://arxiv.org>, 2015.
6. G. Barles and J. Burdeau, *The Dirichlet problem for semilinear second-order degenerate elliptic equations and applications to stochastic exit time control problems*, Comm. Partial Differential Equations **20** (1995), no. 1-2, 129–178.
7. H. Berestycki and P.-L. Lions, *Nonlinear scalar field equations. I. Existence of a ground state*, Arch. Rational Mech. Anal. **82** (1983), no. 4, 313–345.
8. R. S. Cantrell and C. Cosner, *Spatial ecology via reaction-diffusion equations*, Wiley Series in Mathematical and Computational Biology, John Wiley & Sons, Ltd., Chichester, 2003.
9. A. Chaillou and M. Suri, *A posteriori estimation of the linearization error for strongly monotone nonlinear operators*, Journal of Computational and Applied Mathematics **205** (2007), no. 1, 72–87. MR 2324826
10. S. Congreve and T. P. Wihler, *An iterative finite element method for strongly monotone quasi-linear diffusion-reaction problems*, in preparation, 2015.
11. P. Deufhard, *Newtons method for nonlinear problems*, Springer Ser. Comput. Math., 2004.
12. L. Edelstein-Keshet, *Mathematical models in biology*, Classics in Applied Mathematics, vol. 46, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005, Reprint of the 1988 original.
13. L. El Alaoui, A. Ern, and M. Vohralík, *Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems*, Computer Methods in Applied Mechanics and Engineering **200** (2011), no. 37-40, 2782–2795. MR 2811915
14. A. Friedman (ed.), *Tutorials in mathematical biosciences. IV*, Lecture Notes in Mathematics, vol. 1922, Springer, Berlin; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2008, Evolution and ecology, Mathematical Biosciences Subseries.
15. E. M. Garau, P. Morin, and C. Zuppa, *Convergence of an adaptive Kačanov FEM for quasi-linear problems*, Applied Numerical Mathematics. **61** (2011), no. 4, 512–529.
16. W.-M. Ni, *The mathematics of diffusion*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 82, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.
17. A. Okubo and S. A. Levin, *Diffusion and ecological problems: modern perspectives*, second ed., Interdisciplinary Applied Mathematics, vol. 14, Springer-Verlag, New York, 2001.
18. P. H. Rabinowitz, *Minimax methods in critical point theory with applications to differential equations*, CBMS Regional Conference Series in Mathematics, vol. 65, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1986.
19. H.-G. Roos, M. Stynes, and L. Tobiska, *Robust numerical methods for singularly perturbed differential equations*, second ed., Springer Series in Computational Mathematics, vol. 24, Springer-Verlag, Berlin, 2008, Convection-diffusion-reaction and flow problems.
20. J. Smoller, *Shock waves and reaction-diffusion equations*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 258, Springer-Verlag, New York, 1994.
21. W. A. Strauss, *Existence of solitary waves in higher dimensions*, Comm. Math. Phys. **55** (1977), no. 2, 149–162.
22. R. Verfürth, *Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation*, Numer. Math. **78** (1998), no. 3, 479–493.
23. F. Verhulst, *Methods and applications of singular perturbations*, Texts in Applied Mathematics, vol. 50, Springer, New York, 2005, Boundary layers and multiple timescale dynamics.